

A METHOD FOR PROVIDING PERFORMANCE DATA
OF A WEB SERVER

5 Field of Invention

 This invention relates to monitoring the performance status of applications and in particular to a method for providing performance data of a web server.

10 Background of the Invention

15 It is essential to monitor the performance status of applications once they are deployed and, in particular, where such applications are deployed over the Internet or the World-Wide Web ("Web") and where they must provide a certain quality of service. Prior art monitoring programs ("performance tools") typically comprise of three parts: collection, transport, and analysis. The collection may further include a data filtering function to reduce the data collected. The performance data is typically collected into large files for later transport and analysis. An example of such a performance tool is revealed in U.S. Patent No. 5,067,107 entitled "Continuous Computer Performance Measurement Tool That Reduces Operating System Produced Performance Data for Logging into Global, Process, and Workload Files".

20 These prior art performance tools have a number of disadvantages due to propriety standards used by various vendors. The disadvantages include a high cost for developing the transport and analysis parts for each new application to be monitored; a lack of efficiency in that a performance tool can not be used to monitor a number of different applications from different vendors and that
25 an application can not be monitored by a number of different performance tools from different vendors.

There also exists a number of richly functioned performance tools that have already been deployed for performance monitoring of applications. It would be costly to have to re-write and to leverage these performance tools to monitor new applications. However, by divorcing the collection and transport functions from the analysis function, the above disadvantages can be reduced. The "publication" of performance data for analysis by performance tools of a number of different vendors is advantageous. It would be possible to create performance tools that work independently of the applications being monitored. It is therefore an aspect of an object of the present invention for providing "publication" of performance data to such performance tools.

Performance data in this document refers to information that provides a snapshot or a quick scan of what is currently happening on a computing system or server. It includes data like how much memory is being used by what process.

Summary of the Invention

According to an aspect of the invention, there is provided a method for a servlet of an Application server running on a Web server to provide performance data to a performance tool running on a client computer, where the Web server and the client computer are in communication over a network using a network protocol, said method for said servlet comprising: receiving a request for performance data from the performance tool, where the request is transported over the network; obtaining the performance data as per the request; formatting the performance data into a data structure; and providing the data structure for transport to the performance tool over the network.

According to another aspect of the invention, there is provided a monitoring system embedded within a computing system for providing performance data to a performance tool running on a client computer, where the computing system and the client computer are in communication over a network using a network protocol, comprising: a communication system to receive a request

for performance data from the performance tool, where the request is transported over the network; a data collection system to obtain the performance data as per the request; and a processing system to format the performance data into a data structure; wherein the communication system provides the data structure for transport to the performance tool over the network.

According to another aspect of the invention, there is provided an article to provide performance data of a computing system to a performance tool running on a client computer, where the computing system and the client computer are in communication over a network using a network protocol, comprising: a computer-readable storage medium for the computing system; means recorded on the medium for the computing system to receive a request for performance data from the performance tool, where the request is transported over the network; means recorded on the medium to obtain the performance data as per the request; means recorded on the medium to format the performance data into a data structure; and means recorded on the medium to provide the data structure for transport to the performance tool over the network.

According to another aspect of the invention, there is provided means recorded on the medium to provide the data structure for transport to the performance tool over the network.

According to another aspect of the invention, there is provided a method for a performance tool running on a client computer to retrieve performance data from a servlet of an Application server running on a Web server, where the Web server and the client computer are in communication over a network using a network protocol, said method for said performance tool comprising: sending a request for performance data to the servlet, where the request is transported over the network; and receiving a data structure containing the performance data transported from the servlet over the network.

According to another aspect of the invention, there is provided a performance tool embedded within a client computer which retrieves performance data from a servlet of a computing system,

where the computing system and the performance tool are in communication over a network using a network protocol, comprising: a processor to provide a request for performance data where the request is transported to the computing system over the network; and to receive a data structure containing the performance data transported from the computing system over the network.

According to another aspect of the invention, there is provided an article of a performance tool running a client computer for retrieving performance data from a servlet of a computing system, where the computing system and the performance tool are in communication over a network using a network protocol, comprising: a computer-readable storage medium; means recorded on the medium for providing a request for performance data for transport to the computing system over the network; and means recorded on the medium for receiving a data structure containing the performance data transported from the computing system over the network.

Brief Description of the Drawings

In the accompanying drawings:

Figure 1 illustrates an overview of a performance monitoring system in accordance with a preferred embodiment of the present invention; and

Figure 2 illustrates a flowchart outlining the steps for processing a request by a servlet 110 in accordance with Figure 1.

Detailed Description of the Preferred Embodiments

Operating systems run applications. The applications ask the operating system for services, like getting a file. A Web server is an application, or may be a set of applications, running in an operating system providing web services to the Internet. The Web server can run sub-processes independent of the operating system. It does its own control. One of the things it can run is an

Application server. An Application server in turn also runs sub-processes called servlets ("servlets"). Alternatively, a Web server may run servlets.

In a preferred embodiment, a middleware is used to collect performance data for access by other sub-processes through an application program interface ("API"). The middleware is a part of the Application server. The servlet may poke the Application server to obtain the performance data; however, other known means of obtaining performance data are also available.

Alternately, the Web server may keep performance data on the processes that it is running to provide a snapshot of what is currently happening on the server. The performance data is accessible through an API to sub-processes.

The present invention utilizes the existing Internet communication infrastructure for performance tools to provide a facility to transport collected performance data and to provide a standard interface between the collection and the analysis parts. The existing Internet communication infrastructure uses a network protocol known as Transport Control Protocol using Internet Protocol or TCP/IP. However, other embodiments using different network protocols may also be used.

A collection part, such as a servlet, is provided on an Application server to obtain the desired performance data in response to a request from a performance tool running on a client computer. The request is a Hypertext Transfer Protocol ("HTTP") "Get" or "Post". HTTP for the purposes of this document includes HTTP secured. Alternatively, another protocol providing similar functionality can of course be used. The response accordingly provides the performance data in the format of a standard Internet data structure such as Hypertext Markup Language ("HTML") or Extensible Markup Language ("XML") to the requesting program.

The analysis part is preferably implemented as a Java application in a browser as a Java applet, but other techniques such as Extensible Style Language style sheets can also be used for rendering the data. The browser is used to render the performance data as desired. The existing browser rendering and communications functions are leveraged so that these functions do not have to be developed by the performance tool vendors, and thereby reducing the cost of development. The performance tool vendors can instead concentrate on their core function of analyzing performance data. One or more different performance tools can, of course, concurrently request performance data from servlets of an Application server for different purposes.

Referring to the drawings and initially to Figure 1, there is illustrated an overview of a performance monitoring system in accordance with a preferred embodiment of the present invention comprising a client computer 100 having a browser 102 that is running a Java application, a performance tool 100; and a Web server 106 running an Application server 108 having a servlet 110 to obtain performance data from a middleware 112. The performance tool 100 issues HTTP "Get" or "Post" requests. For each request, the servlet 112 obtains the desired performance data and formatting the data as a XML data structure, a XML document, for transport to the performance tool 100 via communication services provided by the Web server 106. The performance data is thus received for rendering and analysis by the performance tool 100.

Referring to Figure 2, there is illustrated a flowchart outlining the steps for processing an HTTP request by the servlet 110 in accordance with Figure 1 of the present invention. Upon receiving and validating the HTTP request, the Web server 106 initiates execution of the servlet 110 if it is not already running. The servlet 110 receives the HTTP request (step 200) and creates a HttpServletResponse Object (step 202). The servlet 110 processes the request i.e. parses the request, (step 204) and determines whether the request is valid or invalid (step 206). If the request is invalid then an error page is written to the HttpServletResponse Object for transport to the performance tool 104 (step 208) and the execution of the servlet ends. If the request is valid then a data structure in XML format, a XML document, is then created (step 210) and the API of the middleware 112 is queried

for performance data (step 212). The obtained performance data is stored in the data structure (step 214). The servlet then determines if further performance observations are required as per the Http request (step 216). If further performance observations for performance data are required then the steps are repeated from step 212, querying for performance data. When no further performance observations are required then the data structure is written to the HttpServletResponse Object (step 208) and the execution of the servlet ends. The HttpServletResponse Object is then passed to the Web server for transport to the requesting performance tool.

Although preferred embodiments of the invention have been described herein, it will be understood by those skilled in the art that variations may be made thereto without departing from the spirit of the invention or the scope of the appended claims.